AUG 22 2005

| REPORT DOCUMENTATION PAGE | Form Approved OMB No. 0704-0188 |
|---|---|



| 1. AGENCY USE ONLY *(Leave blank)* | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | 16.Aug.05 | MAJOR REPORT |

**4. TITLE AND SUBTITLE**
A COMPARISON OF COMORBIDITY MEASUREMENTS TO CONTROL FOR CONFOUNDING IN HEALTH OUTCOMES STUDIES.

**5. FUNDING NUMBERS**

**6. AUTHOR(S)**
CAPT DEVINE JOSHUA W

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
UNIVERSITY OF MINNESOTA MINNEAPOLIS

**8. PERFORMING ORGANIZATION REPORT NUMBER**
CI04-1162

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION AVAILABILITY STATEMENT**
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*

**14. SUBJECT TERMS**

**15. NUMBER OF PAGES**
24

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| | | | |

# A Comparison of Comorbidity Measurements to Control for Confounding in Health Outcomes Studies

JOEL F. FARLEY, B.S. PHARMACY[A]

CAROLYN R. HARLEY, PH.D.[B]

JOSHUA W. DEVINE, PHARM.D., BCPS[A,C]



[A] Ph.D. Candidate, Social & Administrative Pharmacy Graduate Program, College of Pharmacy, University of Minnesota, Minneapolis, MN

[B] Senior Director, Health Economics & Outcomes, i3 Magnify, Eden Prairie, MN

[C] Air Force Institute of Technology, Allied Health Programs, Wright-Patterson AFB, OH

**Address correspondence to:** Joel F. Farley, R.Ph., University of Minnesota, College of Pharmacy, 7-174 Weaver-Densford Hall, 308 Harvard St. S.E., Minneapolis, MN 55455

Email: farl0032@umn.edu, Phone: 612-625-7691, Fax: 612-625-9931

**Article Summary:** This study compares the performance of several comorbidity indices and simple count measurements in the prediction of future health expenditures.

**Funding:** Funding for this study was provided by Pfizer through a dissertation fellowship.

**Manuscript Information:** Including the abstract, tables and references this manuscript is 24 pages in length, contains 27 references, and 4 tables.

**Word Count:** *3498* (Excluding the abstract, references, and tables.)

**Disclaimer:** The views expressed by Captain Joshua Devine in this paper are his alone and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

# A Comparison of Comorbidity Measurements to Control for Confounding in Health Outcomes Studies

**Article Summary:** This study compares the performance of several comorbidity indices and simple count measurements in the prediction of future health expenditures.

**Funding:** Funding for this study was provided by Pfizer through a dissertation fellowship.

**Manuscript Information:** Including the abstract, tables and references this manuscript is 24 pages in length, contains 27 references, and 4 tables.

**Word Count:** *3498* (Excluding the abstract, references, and tables.)

**Objectives:** This study compares the performance of the Elixhauser and Charlson Indices with the Rx-Risk-V score and several simple count measurements including counts of prescriptions, physician's visits, hospital claims, unique prescription classes, and Diagnosis Clusters.

**Study Design:** Simple count measurements, the Elixhauser and Charlson Indices, and the Rx-Risk-V score were calculated one year prior to the filling of a new prescription for an antihypertensive medication for 20,378 members of a managed care organization. The primary outcome variable was the log transformed sum of prescription, physician, and hospital expenditures in the year following the prescription encounter.

**Methods:** In addition to descriptive statistics and spearman correlations between measurements, the predictive performance of each measurement was determined using linear regression models and corresponding $R^2$ statistics.

**Results:** The Charlson Index slightly outperformed the Elixhauser Index ($R^2 = 0.1172$ and 0.1148 respectively) while the prescription claims based Rx-Risk-V ($R^2 = 0.1573$) outperformed both. An age and gender adjusted regression model which included a count of diagnosis clusters was the best individual predictor of payments ($R^2 = 0.1814$). This outperformed age and gender adjusted models of the number of physician's visits ($R^2 = 0.1546$), number of hospital claims ($R^2 = 0.1115$), number of prescriptions filled ($R^2 = 0.1573$), number of unique prescriptions filled ($R^2 = 0.1669$), and log transformed prior health-care payments ($R^2 = 0.1359$).

**Conclusion:** Simple count measurements appear to be better predictors of future expenditures than the comorbidity indices with a count of diagnosis clusters being the single best predictor of future expenditures examined.

**Key Words:** Comorbidity, Risk Adjustment, Prediction, Health Expenditures

## Introduction

Comorbidity scores are a common tool used by researchers in epidemiologic and health services research. Comorbidities are defined as coexisting medical conditions distinct from the primary condition under investigation.[1] Interest in comorbidity scores can be attributed to the importance of relationships between comorbidities and the prognosis, detection, and outcomes of many illnesses.[2] In studies utilizing secondary administrative data, lack of randomization to treatment and control groups can result in health status differences across groups. This can potentially confound the relationship between a treatment and disease under investigation.

Numerous comorbidity controls exist including International Classification of Disease-Ninth Revision-Clinical Modification (ICD-9-CM)-based measures such as the Elixhauser and Charlson Indices[3-7] and pharmacy claims-based measures such as the Chronic Disease Score.[8-12] Simple variables used to measure comorbidity such as counts of medications, physician visits, or medical conditions also have been used in research. These measurements are less complex to implement, and studies have shown them to be as effective, if not more effective, in predicting and controlling for comorbidity.[13-14] Furthermore, these measurements do not suffer from biases related to misclassification. Misclassification can arise in comorbidity indices if complications (conditions arising from the treatment or progression of a condition) are coded as comorbidities (conditions existing simultaneously with and independently of other medical conditions).

Several studies have compared the predictive validity of commonly used comorbidity scores.[13-15] The focus of most of these comparisons, however, has been on the prediction of morbidity and mortality and not on health expenditures. Increasingly, comorbidity scores are being utilized to control for comorbid differences in studies where expenditures and payments are the primary dependent variable.[16-19] Because most

comorbidity measures were developed to predict mortality or morbidity, potential differences may arise when these measures are applied to expenditure outcomes. For example, although patients not surviving the initial onset of an acute myocardial infarction would be included as a death in mortality estimates they may not have significant health expenditures because of their death. In this case, measures that are good predictors of mortality may not necessarily predict expenditures well.

Recently, comorbidity indices have been used to argue for differences in capitation payment rates.[20-21] Of particular concern in capitation arrangements is that a small fraction of insurance beneficiaries often accounts for a large portion of health expenditures. These high expenditure individuals may also be of interest to insurers wishing to target disease management programs to control health plan spending. Although comorbidity indices may be useful in predicting high expenditure individuals, a comparison has not been undertaken to date. Therefore, the use of comorbidity measures for these purposes may be in question.

Given the disparate comorbidity measures currently used and the lack of studies comparing their performance in analyses with expenditures outcomes, this study was undertaken to compare the performance of different comorbidity measures in predicting individual health expenditures. Specifically, the performance of three different administrative claims based comorbidity scores was assessed; the ICD-9-based Charlson and Elixhauser Indices and the pharmacy-claims based Rx-Risk-V Score. Furthermore, the performance of these scores was compared to simpler comorbidity measurements including counts of prescriptions, physician's visits, hospital claims, unique prescription classes, and cumulative number of Diagnosis Clusters serving as a proxy for count of unique health conditions.

## Methods

*Data Source*

For this analysis we used hospital, physician, and pharmacy claims data from a large managed care organization. The study population included 20,378 individuals aged 18 and older who had claims for a diagnosis of hypertension (ICD-9-CM diagnoses 401 – 404, 362.11, or 437.2) and who obtained a new antihypertensive medication between 1/1/2001 to 12/31/2002. The index date in this study was the original prescription purchase date. The pre and post periods encompass the periods one year prior to and one year after the index date respectively. Individuals with gaps in enrollment greater than 31 days during the pre- and post-index periods were excluded.

*Comorbidity Indices*

*Charlson Index*

The Charlson Index is the most common index used currently to control for comorbidity in health outcomes studies. The original Charlson Index was developed for use with medical records and consisted of 19 different diseases weighted according to disease severity as 1, 2, 3, or 6. The index has since been adapted into several 17 item weighted indices for use with administrative data.[1,3,6,22,23] A comprehensive comparison performed by Schneeweiss et al. examined differences in the predictive ability of several Charlson Indices on mortality, long-term care admissions, hospitalizations, physician visits, and expenditures for physician services.[13] Results from this study showed little difference in the performance of different Charlson Indices, with the Romano adaptation performing best. We used a modified version of the Romano adapted Charlson Index to accommodate changes in ICD-9 coding.[23] Based on previous studies which suggest that adding physician claims to hospital claims increases the performance of the Charlson

Index, we ran the index first using pre-period hospital claims and then using both pre-period hospital and physician claims.[24]

*Elixhauser Index*

A relatively new comorbidity measurement is the Elixhauser Index.[5] The Elixhauser Index measures the influence of 30 different comorbid conditions. The index distinguishes comorbidities from complications by considering only secondary diagnoses unrelated to the principal diagnosis through the use of Diagnosis Related Groups (DRGs). For example, a patient with a claim for congestive heart failure would have this condition coded as a comorbidity only if the record did not contain a DRG for cardiac disease. Current coding for the Elixhauser Index was downloaded from the Agency for Healthcare Research and Quality (AHRQ).[25] The index was run first using pre-period hospital claims alone and then using both pre-period hospital and physician claims. Although DRGs are not available within physician claims it was thought that many comorbid conditions would be missed if this data was not included. The final Elixhauser scores were calculated as the sum of comorbid conditions present. Hypertension was excluded from the final score due to the disease population studied.

*Rx-Risk V Score*

A number of indices commonly referred to as Chronic Disease Scores (CDS) have been developed for use with pharmacy claims data.[8,11,12,26] The most recent modifications to the CDS are the Rx-Risk score for use in a general population and the Rx-Risk-V score for use in a Veterans Administration (VA) population.[8,11] Although developed for a VA population, the Rx-Risk-V score was deemed more applicable to our study population based on the population's age and disease distribution. Coding of the Rx-Risk-V was completed using medication classes provided within the original manuscript and corresponding Medi-Span codes.[11] The Rx-Risk-V identifies 45 distinct

comorbid conditions by linking them to medications used in the course of treatment. We used both a non-weighted and weighted count of Rx-Risk-V conditions. Weights for the Rx-Risk-V were taken directly from the original published prospective cost coefficient estimates.[11]

*Simple Counts*

Specific counts of health care utilization in this study included a 12 month count of physician's visits, count of hospital claims, count of prescriptions filled, and count of unique prescriptions (classes of prescriptions) filled in the one year pre-period. As a proxy for the number of unique medical conditions, we placed conditions into diagnosis clusters and used a count of diagnosis clusters as a summary score.[27] Coding of conditions into one of 119 unique diagnosis clusters was performed using ICD-9-CM claims data. Individuals with ICD-9 claims not identified in diagnosis clustering were identified as having an "other" diagnosis cluster. Hypertension was excluded as a cluster based on the population studied. Diagnosis clusters were implemented using hospital claims as well as using both hospital and physician claims. Finally, we used a sum of pre-period payments and sum of log-transformed pre-period payments as potential predictors of post-period payments. To avoid log transformation errors, $0.01 was added to pre-period payments prior to transformation.

*Dependent Variable*

The primary outcome in this study was the sum of individual health care expenditures defined as the sum of hospital, physician, and prescription payments in the one year post-period. Payments included the amount paid out by the health plan to the provider, any amount paid by the patient including deductibles and co-payments, ancillary payments, and any amount reserved from the health plan. Log transformation of the dependent variable was performed to account for non-normality.

*Statistical Analyses*

Descriptive statistics of population characteristics were performed including mean, standard deviation, and range for demographic, payment, health utilization, and comorbidity variables. We assessed correlations between each comorbidity measurement using spearman correlations ($r_s$). Rank-order spearman correlations were used to account for potential non-normality bias in the independent variables. The performance of each comorbidity measurement was assessed through ordinary least squares linear regressions. $R^2$ values representing the amount of variance explained by each regression model were used as the basis for comparisons. Higher $R^2$ values correspond to greater explanation of model variance and thus greater predictive ability. To examine the performance of measurements in predicting high expenditures we employed area under the Receiver Operating Characteristic (ROC) curve comparisons. Area under the ROC comparisons assess the ability of each measurement to accurately predict true positive cases while not predicting true negative cases. Area under the ROC curve outcomes are analyzed through the use of c statistics which can range from 0.5 to 1.0 representing no predictive ability and perfect predictive ability respectively. The ROC outcome was dichotomized 0/1, with 1 representing high expenditure individuals who spent at or above the 90$^{th}$ percentile among the study population.

## Results

*Descriptive Statistics*

Full descriptive statistics for the 20,378 individuals who composed our study population are provided in Table 1. Population members were on average 49 years of age. There were slightly more men (53%) than women. In the one year prior to filling a new prescription for an antihypertensive medication, individuals averaged 1.7 hospital claims, 10.4 physician visits, 13 prescriptions, and used 5 unique prescription

medications. The average count of diagnosis clusters identified through pre-period hospital and physician visit claims was approximately 7 compared to an average of 2 diagnosis clusters identified through hospital claims alone. The average comorbidity score for each diagnosis based index was greater when physician's claims were combined with hospital claims. Average scores ranged from 0.55 for the Charlson Index, 0.62 for the Elixhauser Index, 1.98 for the non-weighed Rx-Risk-V Score, and $4111 for the weighted Rx-Risk-V Score. Compared to average payments of $4615 in the pre-period, payments in the post period averaged $6301. The majority of payments incurred in the post-period were hospital payments ($2756), followed by physician ($2053), and prescription payments ($1492). The cutoff for 90$^{th}$ percentile spending was $12,945. Compared to average post-period spending of $3,3320 among individuals below the 90$^{th}$ percentile cutoff, individuals above the 90$^{th}$ percentile cutoff spent an average of $33,156.

*Correlations*

The strength of correlations between the different indices varied across types of measurements, as seen in Table 2. Correlation between the two ICD-9-claims based indices (the Elixhauser and Charlson) was fair ($r_s = 0.562$) in this analysis. However, correlation between ICD-9- and pharmacy-claims based indices was small with correlations between the non-weighted Rx-Risk-V score and Elixhauser Index being slightly better ($r_s = 0.301$) than between the non-weighted Rx-Risk-V score and Charlson Index ($r_s = 0.242$). Among the count measurements analyzed, there was high correlation between hospital claims identified diagnosis clusters and the number of hospital claims ($r_s = 0.932$) and between hospital and physician's claims identified diagnosis clusters and the number of physician's visits ($r_s = 0.820$). The non-weighted Rx-Risk-V score was strongly correlated with the number of prescriptions used ($r_s = 0.806$) and the number of unique prescriptions used (0.848). Correlations were fair between count of hospital and

physician visits ($r_s = 0.560$) and counts of prescription fills and physician visits ($r_s = 0.555$). However, correlations were small across the count of prescriptions filled and hospital counts ($r_s = 0.327$).

*Predictive Performance*

The predictive performance of each comorbidity measurement on post-period payments is shown in Table 3. The amount of variation explained by each measurement increased without exception when age and gender were included as predictors. Similarly, the addition of physician claims to hospital claims increased the predictive performance of both the Charlson and Elixhauser Indices. The Full Charlson Index which utilized both hospital and physician claims as well as adjustments for age and gender slightly outperformed the Full Elixhauser Index ($R^2 = 0.1172$ and 0.1148 respectively). Among the two Rx-Risk-V scores, the non-weighted score ($R^2 = 0.1381$) was a better predictor of than the weighted score ($R^2 = 0.1261$). Compared to the age and gender included Rx-Risk-V ($R^2 = 0.1573$) both ICD-9 based indices appear slightly inferior. The regression model which included a count of diagnosis clusters, age, and gender was the best individual predictor of future payments ($R^2 = 0.1814$). This outperformed age and gender included models of counts of physician's visits ($R^2 = 0.1546$), hospital claims ($R^2 = 0.1115$), prescriptions filled ($R^2 = 0.1573$), unique prescriptions filled ($R^2 = 0.1669$), and log transformed prior health care payments ($R^2 = 0.1359$).

To examine the effect of adding simple health utilization counts to comorbidity indices, we added simple count controls to the Charlson, Elixhauser, and Rx-Risk-V models. For both of the ICD-9 based indices, the addition of a count of unique prescriptions increased explained variation the greatest. Adding a count of unique prescriptions increased the amount of explained variation 43% in the Charlson Index ($R^2 = 0.2050$) and 42% in the Elixhauser Index ($R^2 = 0.1967$). For the Rx-Risk-V score, the

addition of counts for number of physician visits and hospital visits had the greatest impact on explained variation increasing $R^2$ values 22% ($R^2 = 0.2014$) and 16% ($R^2 = 0.1871$) respectively. The addition of a count of prescriptions to diagnosis clusters was the best combination control for comorbidities in our study ($R^2 = 0.2190$).

Results from the area under the ROC analysis are shown in Table 4. C-statistic values < 0.7, between 0.7 - 0.8, and > 0.8 are generally considered poor, fair, and excellent respectively. None of the c-statistic values in our analysis exceeded 0.7 indicating poor performance in predicting high expenditures. The best performing scores for predicting high expenditures appear to be counts of physician's visits (c = 0.6927), diagnosis clusters (c = 0.6897), and hospital claims (c = 0.6756).

**Discussion**

In the prediction of future expenditures, the Rx-Risk-V score outperformed both the Charlson and Elixhauser Indices. This validates prior comparisons of different Charlson Indices and prescription claims scores in predicting expenditure outcomes.[13,14] Furthermore, it expands upon this research to include the Elixhauser Index as potentially inferior to prescription claims based comorbidity scores in predicting expenditures. Interestingly, in studies examining the performance of ICD-9-based indices to prescription claims-based scores in predicting mortality and morbidity, ICD-9-based scores outperformed prescription claims-based scores. This supports the hypothesis that the impact of comorbidity on expenditure outcomes is different than when applied to morbidity and mortality outcomes.

Several reasons may explain why ICD-9 based indices did not perform as well as the Rx-Risk-V score. First, the Rx-Risk-V score captures more comorbidities than the Elixhauser or Charlson Indices. This could explain why average scores of the Elixhauser (mean = 0.61) and Charlson Indices (mean = 0.55) were lower than the Rx-Risk-V score

(1.98). Second, people are more likely to utilize prescription than hospital or physician services. The average person filled approximately 13 prescriptions in the pre-period compared to an average of 1.7 hospital claims and 10.4 physician's visits. Differences in utilization may also explain why pharmacy claims captured more comorbidities than physician or hospital claims.

This study again suggests that simple measures of utilization are better predictors of future expenditures than elaborate comorbidity indices. With the exception of a count of hospital claims, each simple count measurement examined provided greater explanation to the variance in future health expenditures than either the Elixhauser or Charlson Indices. In addition, a count of prescriptions and count of unique prescriptions provided greater explanatory power than the elaborate Rx-Risk-V score. Schneeweiss et al. showed similar results with the number of distinct medications used during a one year baseline period predicting future expenditures best in a sample of patients 65 years and older.[13] Similarly, Perkins et al. showed that the number of pharmacy subclasses and number of medications used were better predictors of future expenditures than prescription and ICD-9 based comorbidity indices.[14] Our study builds upon these studies by comparing several different count measurements including a proxy for the number of unique medical conditions, diagnosis clusters. A count of diagnosis clusters was the single greatest individual predictor of future health care expenditures in our study.

Utilization measurements provide other advantages besides increased predictive ability. They are easier to implement because they do not require linking claims to specific diagnoses. For example, the Rx-Risk-V requires frequent updating to reflect changes in medication approvals. Similarly, it does not permit classifying a medication into more than one condition leading to potential misclassification. For example, individuals are coded as having liver failure in the Rx-Risk-V if they use lactulose, a

medication commonly used for constipation. One disadvantage in using the Elixhauser Index is that it requires the use of DRGs. In our study, the majority of comorbid conditions identified in the Elixhauser Index were drawn from physician claims which did not have DRG information. This could cause potential misclassification of complications as comorbidities. One disadvantage of comorbidity indices is that they do not capture differences in disease severity. Count measurements may better reflect disease severity by capturing the intensity of resource utilization. As an example, although a person would be identified only as having hypertension in a comorbidity index, a count of the number of medications used to treat hypertension might reflect better the intensity of treatment and potential severity of the condition.

Compared to simple counts of medications and hospital or physician visits, counts of diagnosis clusters and unique prescriptions were better predictors of expenditures. These measurements may be better proxies for the number of medical conditions an individual has than utilization counts. As opposed to a simple count of prescription fills which includes the original fill as well as refills, a count of unique prescriptions more accurately identifies underlying medical conditions. Similarly, a count of diagnosis clusters may better reflect the number of medical conditions than a simple count of physician or hospital visits. This suggests that the number of medical conditions may be a better predictor of future expenditures than past utilization.

The addition of simple count measurements significantly improved the predictive performance of each comorbidity index. For both the Elixhauser and Charlson Indices, the addition of a count of unique prescriptions increased predictive performance better than the addition of counts of physician or hospital claims. For the Rx-Risk-V score, the addition of a count of physician visits increased predictive performance better than the addition of prescription utilization measures. This suggests that the addition of

information from claims sources other than that required in coding the index has the greatest impact on increasing the predictive performance of each index. The marginal benefit of including additional claims information can be weighed against the cost and administration of including additional data sources by individual researchers.

None of the measurements used in this analysis were effective at predicting 90$^{th}$ percentile spending. The best prediction of high expenditures was achieved through simple count measurements. The prior use of physician services was the best predictor of significant expenditures followed by a count of diagnosis clusters and count of hospital claims. Interestingly, prescription claims information was not as accurate in predicting high expenditures as hospital and physician claims information. This may reflect differences in the cost of services. For example, hospital admissions are generally much more expensive than prescription medications. Therefore, individuals using physician and hospital services more frequently will incur higher expenditures.

Our results should be interpreted in light of the following limitations. Because we used claims data for our analyses, information on services not billed to the insurance was unavailable. One potential limitation related to the Elixhauser Index is that DRG information was not available in the physician claims. This may have caused some misclassification of complications as comorbidities. The hospital claims used in this study had 9 diagnosis fields and physician claims had 4 diagnosis fields. This leaves the possibility that individual comorbidities were not identified. In coding the Rx-Risk-V score, some conditions that rely on claims for durable medical equipment such as urinary incontinence and ostomy products were not coded because claims did not exist in the dataset. Finally, caution should be used when generalizing results beyond the study population of continuously enrolled hypertensive patients aged 18 and older from a managed care organization.

## Conclusions

This study builds upon previous comorbidity comparisons to examine the impact of three different indices and several markers of health utilization on the prediction of future health care expenditures. Among the different comorbidity indices examined, the Rx-Risk-V score outperformed both the Charlson and Elixhauser Indices. Our results suggest that a simple count of diagnosis clusters and the number of unique prescriptions used by an individual are the best predictors of future health expenditures. Furthermore, simple count measurements appear better at controlling for the impact of comorbidities than more elaborate comorbidity indices in studies with expenditure outcomes.

1. Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. *J Clin Epidemiol.* 2000;53:1258-67.

2. de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity: A critical review of available methods. *J Clin Epidemiol.* 2003;56:221-9.

3. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol.* 1992;45:613-9.

4. D'Hoore W, Sicotte C, Tilquin C. Risk adjustment in outcome assessment: the Charlson comorbidity index. *Methods Inf Med.*1993;32:382-7.

5. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care.* 1998;36:8-27.

6. Ghali WA, Hall RE, Rosen AK, Ash AS, Moskowitz MA. Searching for an improved clinical comorbidity index for use with ICD-9-CM administrative data. *J Clin Epidemiol.* 1996;49:273-8.

7. Romano PS, Roos LL, Jollis JG. Further evidence concerning the use of a clinical comorbidity index with ICD-9-CM administrative data. *J Clin Epidemiol.* 1993;46:1085-90.

8. Fishman PA, Goodman MJ, Hornbrook MC, Meenan RT, Bachman DJ, O'Keeffe Rosetti MC. Risk adjustment using automated ambulatory pharmacy data: the RxRisk model. *Med Care.* 2003;41:84-99.

9. Gilmer T, Kronick R, Fishman P, Ganiats TG. The Medicaid Rx model: pharmacy-based risk adjustment for public programs. *Med Care.* 2001;39:1188-1202.

10. Johnson RE, Hornbrook MC, Nichols GA. Replicating the chronic disease score (CDS) from automated pharmacy data. *J Clin Epidemiol.* 1994;47:1191-9.

11. Sloan KL, Sales AE, Liu CF, et al. Construction and characteristics of the RxRisk-V: a VA-adapted pharmacy-based case-mix instrument. *Med Care.* 2003;41:761-74.

12. Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *J Clin Epidemiol.* 1992;45:197-203.

13. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. *Am J Epidemiol.* 2001;154:854-64.

14. Perkins AJ, Kroenke K, Unutzer J, et al. Common comorbidity scales were similar in their ability to predict health care costs and mortality. *J Clin Epidemiol.* 2004;57:1040-8.

15. Schneeweiss S, Maclure M. Use of comorbidity scores for control of confounding in studies using administrative databases. *Int J Epidemiol.* 2000;29:891-8.

16. Balkrishnan R, Christensen DB, Bowton DL. Self-reported health status, prophylactic medication use, and healthcare costs in older adults with asthma. *J Am Geriatr Soc.* 2002;50:924-9.

17. Martin BC, Ricci JF, Kotzan JA, Lang K, Menzin J. The net cost of Alzheimer disease and related dementia: a population-based study of Georgia Medicaid recipients. *Alzheimer Dis Assoc Disord.* 2000;14:151-9.

18. McNamara RL, Powe NR, Thiemann DR, Shaffer T, Weller W, Anderson G. Specialty of principal care physician and Medicare expenditures in patients with coronary artery disease: impact of comorbidity and severity. *Am J Manag Care.* 2001;7:261-6.

19. Krop JS, Saudek CD, Weller WE, Powe NR, Shaffer T, Anderson GF. Predicting expenditures for Medicare beneficiaries with diabetes. A prospective cohort study from 1994 to 1996. *Diabetes Care.* 1999;22:1660-6.

**20.** Grasso ME, Weller WE, Shaffer TJ, Diette GB, Anderson GF. Capitation, managed care, and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 1998;158:133-8.

**21.** McNamara RL, Powe NR, Shaffer T, Thiemann D, Weller W, Anderson G. Capitation for cardiologists: accepting risk for coronary artery disease under managed care. *Am J Cardiol.* 1998;82:1178-82.

**22.** D'Hoore W, Bouckaert A, Tilquin C. Practical considerations on the use of the Charlson comorbidity index with administrative data bases. *J Clin Epidemiol.* 1996;49:1429-33.

**23.** Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. *J Clin Epidemiol.* 1993;46:1075-9.

**24.** Zhang JX, Iwashyna TJ, Christakis NA. The performance of different lookback periods and sources of information for Charlson comorbidity adjustment in Medicare claims. *Med Care.* 1999;37:1128-39.

**25.** Comorbidity software: version 3.0. Available at: http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp. Accessed April 28th, 2005.

**26.** Clark DO, Von Korff M, Saunders K, Baluch WM, Simon GE. A chronic disease score with empirically derived weights. *Med Care.* 1995;33:783-95.

**27.** Schneeweiss R, Cherkin DC, Hart LG, et al. Diagnosis clusters adapted for ICD-9-CM and ICHPPC-2. *J Fam Pract.* 1986;22:69-72.

**Table 1: Population Characteristics**

| | Mean | Standard Deviation | Range |
|---|---|---|---|
| ***Pre-Period Sample Characteristics*** | | | |
| Age | 49 | 10.12 | 18 -101 |
| Male – Gender | 0.53 | 0.50 | Male/Female |
| Count of Diagnosis Clusters - Hospital | 2.01 | 2.53 | 0 - 24 |
| Count of Diagnosis Clusters - Hospital + Physician | 7.06 | 4.75 | 0 - 42 |
| Count of Prescriptions Filled | 13.23 | 17.51 | 0 - 268 |
| Count of Unique Prescriptions | 4.66 | 4.71 | 0 - 54 |
| Count of Physician's Visits | 10.42 | 11.88 | 0 - 255 |
| Count of Hospital Claims | 1.72 | 2.75 | 0 - 52 |
| Total Payments in Pre Period | $4,615 | 13,145 | $0 – $508,720 |
| ***Pre-period Index Scoring*** | | | |
| Charlson - Hospital | 0.24 | 0.70 | 0 – 8 |
| Charlson - Hospital + Physician | 0.55 | 1.05 | 0 - 12 |
| Elixhauser – Hospital | 0.21 | 0.62 | 0 - 9 |
| Elixhauser - Hospital + Physician | 0.61 | 0.99 | 0 - 10 |
| Non-weighted RxRisk-V Score | 1.98 | 2.03 | 0 - 14 |
| Weighted RxRisk-V Score | $4,111 | 3,555 | $0 – 34,028 |
| ***Post Period Payment Amount (Dependent Variable)*** | | | |
| Prescription Payments in Post Period | $1,492 | 2,078 | $3 - 46,717 |
| Physician Payments in Post Period | $2,053 | 4,392 | $0 - 237,958 |
| Hospital Payments in Post Period | $2,756 | 13,728 | $0 - 952,772 |
| Total Post Period Payments for non 90% Spenders | $3,320 | 2,858 | $4 - 12,944 |
| Total Post Period Payments for 90% Spenders | $33,156 | 46,294 | $12,952-1,191,433 |
| Total Post Period Payments | $6,301 | 17,362 | $4 - 1,191,433 |

**Table 2: Spearman correlations of comorbidity measurements**

| | C-HL | C-Full | E-HL | E-Full | RR-NW | RR-W | DC-HL | DC-Full | HL | MD | Rx | Unq. Rx | Pre Pay |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Charlson – Hospital (C-HL) | 1 | | | | | | | | | | | | |
| Charlson – Full (C-Full) | 0.652 | 1 | | | | | | | | | | | |
| Elixhauser – Hospital (E-HL) | 0.612 | 0.422 | 1 | | | | | | | | | | |
| Elixhauser – Full (E-Full) | 0.436 | 0.562 | 0.579 | 1 | | | | | | | | | |
| RxRisk-V NonWeighted (RR-NW) | 0.167 | 0.242 | 0.202 | 0.301 | 1 | | | | | | | | |
| RxRisk-V Weighted (RR-W) | 0.198 | 0.281 | 0.221 | 0.310 | 0.898 | 1 | | | | | | | |
| Diagnosis Cluster – HL (DC-HL) | 0.484 | 0.363 | 0.520 | 0.381 | 0.348 | 0.359 | 1 | | | | | | |
| Diagnosis Cluster Full (DC-Full) | 0.314 | 0.402 | 0.349 | 0.491 | 0.543 | 0.518 | 0.625 | 1 | | | | | |

**Table 2 Continued: Spearman correlations of comorbidity measurements**

| | C - HL | C- Full | E- HL | E- Full | RR- NW | RR- W | DC- HL | DC- Full | HL | MD | Rx | Unq. Rx | Pre Pay |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Hospital Count (HL) | 0.414 | 0.317 | 0.416 | 0.324 | 0.351 | 0.363 | 0.932 | 0.587 | 1 | | | | |
| Physician Count (MD) | 0.283 | 0.364 | 0.300 | 0.415 | 0.538 | 0.514 | 0.544 | 0.820 | 0.560 | 1 | | | |
| Prescription Count (RX) | 0.156 | 0.249 | 0.188 | 0.303 | 0.806 | 0.726 | 0.316 | 0.544 | 0.327 | 0.555 | 1 | | |
| Unique Rx Count (Unq. Rx) | 0.165 | 0.249 | 0.202 | 0.305 | 0.848 | 0.793 | 0.367 | 0.614 | 0.376 | 0.601 | 0.897 | 1 | |
| Ln Pre-ayment (Pre Pay) | 0.420 | 0.436 | 0.420 | 0.445 | 0.564 | 0.532 | 0.705 | 0.742 | 0.674 | 0.774 | 0.597 | 0.596 | 1 |

**Table 3: Percent variation explained by each comorbidity control on log-transformed health expenditures**

| Predictor | Adjusted $R^2$ |
|---|---|
| ***Charlson Index*** | |
| Charlson - Hospital Claims Only | 0.0761 |
| Charlson - Doctor and Hospital Claims | 0.0971 |
| Full Charlson -Including age & gender | 0.1172 |
| Full Charlson + Count of Physician's Visits | 0.1830 |
| Full Charlson + Count of Hospital Claims | 0.1524 |
| Full Charlson + Count of Prescriptions | 0.1999 |
| Full Charlson + Count of Unique Prescriptions | 0.2050 |
| ***Elixhauser Index*** | |
| Elixhauser - Hospital Claims Only | 0.0666 |
| Elixhauser - Doctor and Hospital Claims | 0.0934 |
| Full Elixhauser - Including age & gender | 0.1148 |
| Full Elixhauser + Count of Physician's Visits | 0.1757 |
| Full Elixhauser + Count of Hospital Claims | 0.1504 |
| Full Elixhauser + Count of Prescriptions | 0.1911 |
| Full Elixhauser + Count of Unique Prescriptions | 0.1967 |
| ***RxRisk-V Score*** | |
| RxRisk-V Weighted Score | 0.1261 |
| RxRisk-V Non-Weighted Score | 0.1381 |
| RxRisk-V Score Non-Weighted - Including age and gender | 0.1573 |
| RxRisk-V Score Non-Weighted + Count of Physician's Visits | 0.2014 |
| RxRisk-V Score Non-Weighted + Count of Hospital Claims | 0.1871 |
| RxRisk-V Score Non-Weighted + Count of Prescriptions | 0.1756 |
| RxRisk-V Score Non-Weighted + Count of Unique Prescriptions | 0.1739 |

**Table 3 Continued: Percent variation explained by each comorbidity control on log-transformed health expenditures**

| Predictor | Adjusted $R^2$ |
|---|---|
| ***Diagnosis Clustering*** | |
| Diagnosis Clustering - Hospital Claims Only | 0.1149 |
| Full Diagnosis Clustering - Doctor and Hospital Claims | 0.1711 |
| Full Diagnosis Clustering with age and gender | 0.1814 |
| Full Diagnosis Clustering + Count of Physician's Visits | 0.1969 |
| Full Diagnosis Clustering + Count of Hospital Claims | 0.1902 |
| Full Diagnosis Clustering + Count of Prescriptions | 0.2190 |
| Full Diagnosis Clustering + Count of Unique Prescriptions | 0.2077 |
| ***Simple Counts*** | |
| Count of Physician's Visits | 0.1354 |
| Count of Physician's Visits with age and gender | 0.1546 |
| Count of Hospital Claims | 0.0883 |
| Count of Hospital Claims with age and gender | 0.1115 |
| Count of Prescriptions | 0.1432 |
| Count of Prescriptions with age and gender | 0.1573 |
| Count of Unique Prescriptions (Prescription Classes) | 0.1449 |
| Count of Unique Prescriptions (Prescription Classes) with age and gender | 0.1669 |
| Sum Payments in year prior to index date | 0.1031 |
| Sum Payments in year prior to index date with age and gender | 0.1046 |
| Log transformed sum of payments[1] in year prior to index date | 0.1161 |
| Log transformed sum of payments[1] in year prior to index date with age and gender | 0.1359 |

**Table 3 Continued: Percent variation explained by each comorbidity control on log-transformed health expenditures**

| Predictor | Adjusted $R^2$ |
|---|---|
| ***Full Models*** | |
| Count of Physician, Hospital, & Unique Prescriptions | 0.2087 |
| Full Charlson + Count of Physician, Hospital, & Unique Prescriptions | 0.2263 |
| Full Elixhauser + Count of Physician, Hospital, & Unique Prescriptions | 0.2196 |
| RxRisk-V + Count of Physician, Hospital, & Unique Prescriptions | 0.2140 |
| Diagnosis Clusters + Count of Physician, Hospital, & Unique Prescriptions | 0.2203 |

***1 = $0.01 added to pre-payments prior to log transformation to avoid censoring at $0.00***

**Table 4: Area under the receiver operating characteristic curve predictions of individuals spending at the 90% level**

| Predictor | Area Under the ROC |
|---|---|
| ***Charlson Index*** | |
| Charlson - Hospital Claims Only | 0.6277 (0.6210, 0.6343) |
| Charlson - Doctor and Hospital Claims | 0.6601 (0.6535, 0.6666) |
| ***Elixhauser Index*** | |
| Elixhauser - Hospital Claims Only | 0.6181 (0.6114, 0.6248) |
| Elixhauser - Doctor and Hospital Claims | 0.6555 (0.6489, 0.6620) |
| ***RxRisk-V Score*** | |
| RxRisk-V Score | 0.6412 (0.6345, 0.6478) |
| ***Diagnosis Clustering*** | |
| Diagnosis Clustering - Hospital Claims Only | 0.6872 (0.6808, 0.6935) |
| Full Diagnosis Clustering - Doctor and Hospital Claims | 0.6897 (0.6833, 0.6960) |
| ***Simple Counts*** | |
| Count of Physician's Visits | 0.6927 (0.6863, 0.6990) |
| Count of Hospital Claims | 0.6756 (0.6691, 0.6820) |
| Count of Prescriptions | 0.6357 (0.6290, 0.6423) |
| Count of Unique Prescriptions (Prescription Classes) | 0.6442 (0.6376, 0.6508) |